

<110> Bristol-Myers Squibb Company Chiang, Shu-Jen Jonathan, Basch

<120> DIRECT PRODUCTION OF DESACETYLCEPHALOSPORIN C

<130> ON0163

<140> 09/801,852

<141> 2001-03-08

<150> 60/188,033

<151> 2000-03-09

<160> 15

<170> PatentIn version 3.0

<210> 1

<211> 1716

<212> DNA

<213> Rhodosporidium toruloides

<400> 1

atgeteetta acetetteae eetegeetee etegetgega egeteeaget egeetttgee 60 tetecgacet ecetegteeg eegeacgaac ecaaacgage eceeteeegt egtegacete 120 ggctacgccc gctaccaagg ctacttgaac gagaccgccg gactctactg gtggcgcgga 180 atcegetacg ceteggetea gegettecag getecteaga egecegegae geacaaggee 240 gtccgcaacg cgactgagta tggaccgatc tgttggccgg ctagcgaggg aaccaacacg 300 accaaggget tgccgccgcc tagcaacagc tcgagcagcg cgccgcagaa acaggcgtcg 360 gaggattgcc tetteeteaa tgtegttgcc eeegeegget egtgegaggg egacaatett 420 eccgteeteg tetacattea eggaggtgge taegeetteg gegatgegag caeeggeage 480 gactttgccg ccttcaccaa gcacacggga accaagatgg tcgttgtaaa tctccagtac 540 cgtctcggca gctttggttt cctcgctggc caagccatga aggactacgg tgtaacgaac 600 gccggcttgc ttgaccagca attcgccctt caatgggttc aacagcacgt ctcgaagttc 660 ggcggcaacc ccgatcacgt tacgatttgg ggcgagtctg caggcgcagg gtccgttatg 720 aaccagatca ttgcgaacgg cggcaacacc gtcaaggctc tcggtctcaa gaagcccctc 780 ttccacgctg ccatcggctc ctccgtcttc ctcccctacc aagccaagta caactccccc 840 ttcgccgagc tgctctactc ccaactcgtc tcggcgacaa actgcaccaa agccgcctcg 900 teettegett geetegaage tgtegaeget geggegeteg etgeggeggg egtgaagaae 960 teggeggegt tecegttegg gttttggteg tatqteeeqq tegteqaegq qaeettettq 1020 actgagegeg egtegeteet tetegeeaag ggeaagaaga aceteaatgg caacetette 1080 accgggatca acaacctcga cgaaggattc atattcactg acgccactat tcagaacgac 1140 acgatcageg accagtegea gegegtetee cagttegace geeteetege eggeetette 1200 ccctacatca cctcggagga gcgccaggcc gtcgcgaagc agtacccgat ctccgacqcq 1260 ccgtcaaagg gcaacacett etetegeate teggeegtea tegeggaete gaeettegte 1320 tgcccgacct actggaccgc cgaggcgttc ggctcgtccg cccacaaggg cctcttcgac 1380 tacgegeegg eteaceaege gacegaeaae tegtaetaea teggeteeat etggaaegge 1440 aagaagtogg totogtoogt coagtootto gacggogogo toggoggott catogagaog 1500 ttcaacccga acaacaacgc tgccaacaag accatcaacc cttactggcc gacgttcgac 1560 tegggeaage ageteetett caacaegaeg aegagggaea eeetetetee egeegaeeeg 1620 cgcatcgttg agacttcaag cttgaccgac tttggcacga gccagaagac caagtgcgac 1680 ttctggcgtg ggtcaatctc ggtgaacgcg ggtctc 1716

<210> 2

<211> 572

<212> PRT

<213> Rhodosporidium toruloides

<400> 2

Met Leu Leu Asn Leu Phe Thr Leu Ala Ser Leu Ala Ala Thr Leu Gln 1 5 10 15

Leu Ala Phe Ala Ser Pro Thr Ser Leu Val Arg Arg Thr Asn Pro Asn 20 25 30

Glu Pro Pro Pro Val Val Asp Leu Gly Tyr Ala Arg Tyr Gln Gly Tyr 35 40 45

Leu Asn Glu Thr Ala Gly Leu Tyr Trp Trp Arg Gly Ile Arg Tyr Ala 50 55 60

Ser Ala Gln Arg Phe Gln Ala Pro Gln Thr Pro Ala Thr His Lys Ala 65 70 75 80

Val Arg Asn Ala Thr Glu Tyr Gly Pro Ile Cys Trp Pro Ala Ser Glu 85 90 95

Gly Thr Asn Thr Thr Lys Gly Leu Pro Pro Pro Ser Asn Ser Ser Ser 100 105 110

Ser Ala Pro Gln Lys Gln Ala Ser Glu Asp Cys Leu Phe Leu Asn Val 115 Val Ala Pro Ala Gly Ser Cys Glu Gly Asp Asn Leu Pro Val Leu Val 130 Tyr Ile His Gly Gly Gly Tyr Ala Phe Gly Asp Ala Ser Thr Gly Ser 145 Phe Ala Ala Phe Thr Lys His Thr Gly Thr Lys Met Val Val 175 Val

Asn Leu Gln Tyr Arg Leu Gly Ser Phe Gly Phe Leu Ala Gly Gln Ala 180 185 190

Met Lys Asp Tyr Gly Val Thr Asn Ala Gly Leu Leu Asp Gln Gln Phe 195 200 205

Ala Leu Gln Trp Val Gln Gln His Val Ser Lys Phe Gly Gly Asn Pro 210 215 220

Asp His Val Thr Ile Trp Gly Glu Ser Ala Gly Ala Gly Ser Val Met 225 230 235 240

Asn Gln Ile Ile Ala Asn Gly Gly Asn Thr Val Lys Ala Leu Gly Leu 245 250 . 255

Lys Lys Pro Leu Phe His Ala Ala Ile Gly Ser Ser Val Phe Leu Pro 260 265 270

Tyr Gln Ala Lys Tyr Asn Ser Pro Phe Ala Glu Leu Leu Tyr Ser Gln 275 280 285

Leu Val Ser Ala Thr Asn Cys Thr Lys Ala Ala Ser Ser Phe Ala Cys 290 295 300

Leu Glu Ala Val Asp Ala Ala Ala Leu Ala Ala Ala Gly Val Lys Asn 305 310 315 320

Ser Ala Ala Phe Pro Phe Gly Phe Trp Ser Tyr Val Pro Val Val Asp 325 330 335

Gly Thr Phe Leu Thr Glu Arg Ala Ser Leu Leu Leu Ala Lys Gly Lys 340 345 350

Lys Asn Leu Asn Gly Asn Leu Phe Thr Gly Ile Asn Asn Leu Asp Glu 355 360 365

Gly Phe Ile Phe Thr Asp Ala Thr Ile Gln Asn Asp Thr Ile Ser Asp 370 375 380

Gln Ser Gln Arg Val Ser Gln Phe Asp Arg Leu Leu Ala Gly Leu Phe 385 390 395 400

Pro Tyr Ile Thr Ser Glu Glu Arg Gln Ala Val Ala Lys Gln Tyr Pro 405 410 415

	Ile	Ser	Asp	Ala 420	Pro	Ser	Lys	Gly	Asn 425	Thr	Phe	Ser	Arg	Ile 430	Ser	Ala		
	Val	Ile	Ala 435	Asp	Ser	Thr	Phe	Val 440	Cys	Pro	Thr	Tyr	Trp 445	Thr	Ala	Glu		
	Ala	Phe 450	Gly	Ser	Ser	Ala	His 455	Lys	Gly	Leu	Phe	Asp 460	Tyr	Ala	Pro	Ala		
	His 465	His	Ala	Thr	Asp	Asn 470	Ser	Tyr	Tyr	Ile	Gly 475	Ser	Ile	Trp	Asn	Gly 480		
	Lys	Lys	Ser	Val	Ser 485	Ser	Val	Gln	Ser	Phe 490	Asp	Gly	Ala	Leu	Gly 495	Gly	•	
	Phe	Ile	Glu	Thr 500	Phe	Asn	Pro	Asn	Asn 505	Asn	Ala	Ala	Asn	Lys 510	Thr	Ile		
	Asn	Pro	Tyr 515	Trp	Pro	Thr	Phe	Asp 520	Ser	Gly	Lys	Gln	Leu 525	Leu	Phe	Asn		
	Thr	Thr 530	Thr	Arg	Asp	Thr	Leu 535	Ser	Pro	Ala	Asp	Pro 540	Arg	Ile	Val	Glu		•
	Thr 545	Ser	Ser	Leu	Thr	Asp 550	Phe	Gly	Thr	Ser	Gln 555	Lys	Thr	Lys	Cys	Asp 560		
	Phe	Trp	Arg	Gly	Ser 565	Ile	Ser	Val	Asn	Ala 570	Gly	Leu						
<210> 3 <211> 2220 <212> DNA <213> Rhodosporidium toruloides																		
<400> 3																		
	ggat	ccac	cc c	gaact	ctgt	c co	gctt	tctc	gct	ttct	tcc	ttg	ctgto	ege (	ccat	cgcct		60
	ttcc	cgac	etc c	geege	cato	jc to	ctta	acct	ctt	cacc	cctc	gcct	ccct	cg (	ctgc	gacgct		120
	ccaç	ctcg	jcc t	ttgc	ctct	c cg	gacct	ccct	cgt	ccgc	ecgc	acga	acco	caa a	acgaç	geeeee		180
	tccc	gtcg	ıtc <u>c</u>	gacct	cggc	t ac	gccc	gcta	cca	aggo	ctac	ttga	acga	aga (	ccgc	eggact		240
	ctac	tggt	gg c	gegg	gaato	c go	ctaco	cctc	ggc	tcag	gcgc	ttco	caggo	ctc o	ctcaç	gacgcc		300
	cgcg	acgo	ac a	aggc	cgto	c go	aacç	ıcgac	: tga	igtat	gga	ccga	atct	gtt (	ggcc	gctag		360
	cgag	ggaa	cc a	acac	gaco	a aç	ggct	tgcc	gcc	gcct	agc	aaca	agcto	ga (	gcago	gegee		420
	gcag	aaac	ag ç	gcgtc	ggag	ıg at	tgcc	tctt	cct	caat	gtc	gtt	geee	ccg (	ccgg	ctcgtg		480
	cgag	ggcg	ac a	atct	tccc	g to	ctcç	ıtcta	cat	tcac	gga	ggt	gcta	acg d	cctt	eggega		540
	tgcg	agca	cc g	gcag	cgac	t tt	gccg	cctt	cac	caag	jcac	acgo	ggaad	cca a	agato	gtcgt		600

tgtaaatctc cagtaccgtc tcggcagctt tggtttcctc gctggccaag ccatgaagga

ctacggtgta	acgaacgccg	gcttgcttga	ccaggtgagt	ttcccgcatg	atacccgccc	720
acctttcgac	tcatgctgac	gcctctcccg	ctcgcagcaa	ttcgcccttc	aatgggttca	780
acagcacgtc	tcgaagttcg	gcggcaaccc	cgatcacgtt	acgatttggg	gcgagtctgc	840
aggcgcaggg	tccgttatga	accagatcat	tgcgaacgtg	agccacccga	accgatctcc	900
agccgacttt	cccccccc	cccccccgc	tgacctccct	cgtcttgcag	ggcggcaaca	960
ccgtcaaggc	tctcggtctc	aagaagcccc	tcttccacgc	tgccatcggc	tcctccgtct	1020
tcctccccta	ccaagccaag	tacaactccc	ccttcgccga	gctgctctac	tcccaactcg	1080
tctcggcgac	aaactgcacc	aaagccgcct	cgtccttcgc	ttgcctcgaa	gctgtcgacg	1140
ctgcggcgct	cgctgcggcg	ggcgtgaaga	actcggcggc	gttcccgttc	gggttttggt	1200
cgtatgtccc	ggtcgtcgac	gggaccttct	tgactgagcg	cgcgtcgctc	cttctcgcca	1260
agggcaagaa	gaacctcaat	ggcgtgcgtg	gcgagctttc	gagtgcttca	ggatctcgct	1320
gacactgtcg	accggctcgc	agaacctctt	caccgggatc	aacaacctcg	acgaagatga	1380
gttcccgtcg	acggctctgt	tegeceageg	agactgactt	gttcttttgc	gaagattacg	1440
attcatattc	actgacgcca	ctattcagaa	cgacacgatc	agcgaccagt	cgcagcgcgt	1500
ctcccagttc	gaccgcctcc	tegeeggeet	cttcccctac	atcacctcgg	aggagcgcca	1560
ggccgtcgcg	aagcagtacc	cgatctccga	cgcgccgtca	aagggcaaca	ccttctctcg	1620
catctcggcc	gtcatcgcgg	actcgacctt	cgtgtgcgtt	ccccgtcgtc	ttctccgagt	1680
attccgctga	cttcccgctt	gcccgcagct	gcccgaccta	ctggaccgcc	gaggcgttcg	1740
gctcgtccgc	ccacaagggc	ctcttcgact	acgcgccggc	tcaccacgcg	accgacaact	1800
cgtactacat	cggctccatc	tggaacggca	agaagtcggt	ctcgtccgtc	cagtccttcg	1860
acggcgcgct	cggcggcttc	atcgagacgt	tcaacccgaa	caacaacgct	gccaacaaga	1920
ccatcaaccc	ttactggccg	acgttcgact	cgggcaagca	gctcctcttc	aacacgacga	1980
cgagggacac	cctctctccc	gccgacccgc	gcatcgttga	gacttcaagc	ttgaccgact	2040
ttggcacgag	ccagaagacc	aagtgcgact	tctggcgtgg	gtcaatctcg	gtgaacgcgg	2100
gtctctaggc	gtctttcctt	ccgacttcct	tcgttctttc	gttgtttatt	cttgcagttc	2160
cgttgtatcg	gccattcgtg	cgtgtagctc	actcgagtat	agacgttggc	aagtgcgaaa	2220

<sup>&</sup>lt;210> 4 <211> 544 <212> PRT

## <213> Rhodosporidium toruloides

<400> 4

Thr Asn Pro Asn Glu Pro Pro Pro Val Val Asp Leu Gly Tyr Ala Arg 1 5 10 15

Tyr Gln Gly Tyr Leu Asn Glu Thr Ala Gly Leu Tyr Trp Trp Arg Gly
20 25 30

Ile Arg Tyr Ala Ser Ala Gln Arg Phe Gln Ala Pro Gln Thr Pro Ala
35 40 45

Thr His Lys Ala Val Arg Asn Ala Thr Glu Tyr Gly Pro Ile Cys Trp 50 55 60

Pro Ala Ser Glu Gly Thr Asn Thr Thr Lys Gly Leu Pro Pro Pro Ser 65 70 75 80

Asn Ser Ser Ser Ser Ala Pro Gln Lys Gln Ala Ser Glu Asp Cys Leu 85 90 95

Phe Leu Asn Val Val Ala Pro Ala Gly Ser Cys Glu Gly Asp Asn Leu 100 105 110

Pro Val Leu Val Tyr Ile His Gly Gly Gly Tyr Ala Phe Gly Asp Ala 115 120 125

Ser Thr Gly Ser Asp Phe Ala Ala Phe Thr Lys His Thr Gly Thr Lys 130 135 140

Met Val Val Asn Leu Gln Tyr Arg Leu Gly Ser Phe Gly Phe Leu 145 150 155 160

Ala Gly Gln Ala Met Lys Asp Tyr Gly Val Thr Asn Ala Gly Leu Leu 165 170 175

Asp Gln Gln Phe Ala Leu Gln Trp Val Gln Gln His Val Ser Lys Phe 180 185 190

Gly Gly Asn Pro Asp His Val Thr Ile Trp Gly Glu Ser Ala Gly Ala 195 200 205

Gly Ser Val Met Asn Gln Ile Ile Ala Asn Gly Gly Asn Thr Val Lys 210 215 220

Ala Leu Gly Leu Lys Lys Pro Leu Phe His Ala Ala Ile Gly Ser Ser 225 230 235 240

Val Phe Leu Pro Tyr Gln Ala Lys Tyr Asn Ser Pro Phe Ala Glu Leu 245 250 255

Leu Tyr Ser Gln Leu Val Ser Ala Thr Asn Cys Thr Lys Ala Ala Ser 260 265 270

Ser Phe Ala Cys Leu Glu Ala Val Asp Ala Ala Ala Leu Ala Ala Ala 275 280 285

Gly Val Lys Asn Ser Ala Ala Phe Pro Phe Gly Phe Trp Ser Tyr Val 295 Pro Val Val Asp Gly Thr Phe Leu Thr Glu Arg Ala Ser Leu Leu Leu 310 315 Ala Lys Gly Lys Lys Asn Leu Asn Gly Asn Leu Phe Thr Gly Ile Asn 330 Asn Leu Asp Glu Gly Phe Ile Phe Thr Asp Ala Thr Ile Gln Asn Asp 345 Thr Ile Ser Asp Gln Ser Gln Arg Val Ser Gln Phe Asp Arg Leu Leu 360 . 365 Ala Gly Leu Phe Pro Tyr Ile Thr Ser Glu Glu Arg Gln Ala Val Ala 380 Lys Gln Tyr Pro Ile Ser Asp Ala Pro Ser Lys Gly Asn Thr Phe Ser 390 395 Arg Ile Ser Ala Val Ile Ala Asp Ser Thr Phe Val Cys Pro Thr Tyr 405 Trp Thr Ala Glu Ala Phe Gly Ser Ser Ala His Lys Gly Leu Phe Asp 425 Tyr Ala Pro Ala His His Ala Thr Asp Asn Ser Tyr Tyr Ile Gly Ser Ile Trp Asn Gly Lys Lys Ser Val Ser Ser Val Gln Ser Phe Asp Gly Ala Leu Gly Gly Phe Ile Glu Thr Phe Asn Pro Asn Asn Asn Ala Ala 470 Asn Lys Thr Ile Asn Pro Tyr Trp Pro Thr Phe Asp Ser Gly Lys Gln 485 Leu Leu Phe Asn Thr Thr Thr Arg Asp Thr Leu Ser Pro Ala Asp Pro 505 Arg Ile Val Glu Thr Ser Ser Leu Thr Asp Phe Gly Thr Ser Gln Lys 520 Thr Lys Cys Asp Phe Trp His Gly Ser Ile Ser Val Asn Ala Gly Leu 530 535

<210> 5
<211> 15
<212> PRT
<213> Rhodosporidium toruloides

<400> 5

```
Thr Asn Pro Asn Glu Pro Pro Pro Val Val Asp Leu Gly Tyr Ala
    <210> 6
    <211> 24
    <212> DNA
    <213> Other nucleic acid
    <400> 6
                                                                          24
    gatcacccgg gttgggccca ctag
    <210> 7
    <211> 30
    <212> DNA
    <213> Other nucleic acid
    <400> 7
                                                                          30
    actcgccgcc atggtcctta acctcttcac
    <210> 8
ij
    <211> 30
    <212> DNA
    <213> Other nucleic acid
ij,
    <400> 8
    gaaagacccc tagagacccg cgttcaccga
                                                                          30
    <210> 9
ing i
    <211>
N
    <212> PRT
g:
    <213> Rhodosporidium toruloides
C
    <400> 9
    Thr Asn Pro Asn Glu Pro
                    5
    <210> 10
    <211> 17
    <212> PRT
    <213> Other nucleic acid
    <400> 10
    Ala Cys Asn Ala Ala Tyr Cys Cys Asn Ala Ala Tyr Gly Ala Arg Cys
    Cys
    <210> 11
    <211> 17
    <212> PRT
    <213> Other nucleic acid
```

```
<400> 11
Gly Gly Tyr Thr Cys Arg Thr Thr Asn Gly Gly Arg Thr Thr Asn Gly
Thr
<210> 12
<211> 17
<212> PRT
<213> Other Nucleic Acid
<400> 12
Gly Gly Tyr Thr Cys Arg Thr Thr Gly Gly Gly Arg Thr Thr Asn Gly
Thr
<210> 13
<211> 17
<212> PRT
<213> Other nucleic acid
<400> 13
Gly Gly Tyr Thr Cys Arg Thr Thr Ala Gly Gly Arg Thr Thr Asn Gly
Thr
<210> 14
<211>
      17
<212>
      PRT
<213> Other nucleic acid
<400> 14
Gly Gly Tyr Thr Cys Arg Thr Thr Thr Gly Gly Arg Thr Thr Asn Gly
                                   10
Thr
<210>
      15
<211> 17
<212> PRT
<213> Other nucleic acid
<400> 15
```

ļ,

Q)

The state

T. I.J. C.

Gly Gly Tyr Thr Cys Arg Thr Thr Cys Gly Gly Arg Thr Thr Asn Gly 1 5 10 15

Thr